Risk Analysis, Vol. 40, No. 1, 2020 DOI: 10.1111/risa.13004

An Integrated Scenario Ensemble-Based Framework for Hurricane Evacuation Modeling: Part 2—Hazard Modeling

Brian Blanton,[1,*] Kendra Dresback,[2] Brian Colle,[3] Randy Kolar,[2] Humberto Vergara,[4] Yang Hong,[4] Nicholas Leonardo,[3] Rachel Davidson,[5] Linda Nozick,[6] and Tricia Wachtendorf[7]

Hurricane track and intensity can change rapidly in unexpected ways, thus making predictions of hurricanes and related hazards uncertain. This inherent uncertainty often translates into suboptimal decision-making outcomes, such as unnecessary evacuation. Representing this uncertainty is thus critical in evacuation planning and related activities. We describe a physics-based hazard modeling approach that (1) dynamically accounts for the physical interactions among hazard components and (2) captures hurricane evolution uncertainty using an ensemble method. This loosely coupled model system provides a framework for probabilistic water inundation and wind speed levels for a new, risk-based approach to evacuation modeling, described in a companion article in this issue. It combines the Weather Research and Forecasting (WRF) meteorological model, the Coupled Routing and Excess STorage (CREST) hydrologic model, and the ADvanced CIRCulation (ADCIRC) storm surge, tide, and wind-wave model to compute inundation levels and wind speeds for an ensemble of hurricane predictions. Perturbations to WRF's initial and boundary conditions and different model physics/parameterizations generate an ensemble of storm solutions, which are then used to drive the coupled hydrologic + hydrodynamic models. Hurricane Isabel (2003) is used as a case study to illustrate the ensemble-based approach. The inundation, river runoff, and wind hazard results are strongly dependent on the accuracy of the mesoscale meteorological simulations, which improves with decreasing lead time to hurricane landfall. The ensemble envelope brackets the observed behavior while providing "best-case" and "worst-case" scenarios for the subsequent risk-based evacuation model.

KEY WORDS: Coupled models; hurricane; river flow; storm surge; uncertainty

1. INTRODUCTION

Coastal flooding and high winds during hurricane events can cause substantial damage and hinder evacuation of threatened coastal areas. Well before coastal impacts are experienced, the uncertainty in hurricane trajectory and intensity poses challenges

[1]Renaissance Computing Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.
[2]School of Civil Engineering and Environmental Science, University of Oklahoma, Norman, OK, USA.
[3]School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY, USA.
[4]Hydrometeorology and Remote Sensing Laboratory, University of Oklahoma, Norman, OK, USA.
[5]Department of Civil and Environmental Engineering, University of Delaware, Newark, DE, USA.
[6]Civil and Environmental Engineering, Cornell University, Ithaca, NY, USA.
[7]Department of Sociology and Criminal Justice, University of Delaware, Newark, DE, USA.
*Address correspondence to Brian Blanton, Renaissance Computing Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27517, USA; Brian_Blanton@Renci.Org.

0272-4332/20/0100-0117$22.00/1

for decisionmakers and emergency managers who need to plan evacuation activities. Hazard modeling for threatened coastal areas is thus a critical part of the evacuation planning and decision making. However, most official evacuation planning studies use relatively simplified models for hurricane winds and related factors, such as those provided by (HURREVAC, 2016), and storm surge threat, as provided by the NOAA Sea, Lake, and Overland Surges from Hurricanes (SLOSH) (Jelesnianski, Chen, & Shaffer, 1992) model, and they do not include precipitation and inland waters routing downstream toward coastal regions. Additionally, there is little (if any) accounting for uncertainty in storm evolution and impacts, which is usually relatively large, and thus evacuation studies do not easily focus on the risk involved in moving and/or sheltering the population under threat. Considering that evacuation costs are generally substantial (Whitehead, 2003), the need to skillfully simulate potential storm surge, winds, runoff, and waves is essential for minimizing public costs to (potentially) avoid unnecessary evacuations.

The companion article by Davidson et al. (2020) introduces a new integrated, risk-based evacuation order decision support framework that captures the dynamic interactions among the natural, human, and infrastructure systems. The evacuation model within the framework is a multistage stochastic program that minimizes the risk for a threatened population, as well as its total travel time, and is a logical extension of Li, Xu, Nozick, and Davidson (2011) and Apivatanagul, Davidson, and Nozick (2012). The framework offers a substantial advance in the state of the art by providing a well-hedged solution that is robust under the range of ways the hurricane might evolve, and that takes advantage of the substantial value of increasing information (or decreasing degree of uncertainty) over the course of a hurricane event. This new approach to evacuation order decision support would be impossible without a representation of the hazard that explicitly captures the uncertainty in the hurricane's development over time, a description that did not previously exist. This article focuses on the hazard modeling portion of the framework, the part that provides that description. Specifically, the hazard modeling system in this article captures uncertainty in the hurricane's life cycle through an ensemble of meteorological model perturbations (each member of which is also referred to as a scenario), which, in turn, impacts the overland, precipitation-driven flooding, the downstream routing of inland flooding to the coastal ocean, and the coastal storm surge. It integrates three state-of-the-art, physics-based numerical models for mesoscale meteorology, inland rainfall accumulation and routing, and coastal storm surge, waves, and tides. North Carolina (NC) coastal counties serve as a test bed for implementation of the framework. To our knowledge, this comprehensive approach to modeling the hazard for evacuation modeling and studies is not used. Typically, catalogs of SLOSH storm surge simulations and traffic clearance times are used to define evacuation zones, without consideration of the potential contributions from riverine inflows to the coastal zone, fully dynamic tides, and, more critically, probabilistic assessment of how these hazards evolve as the event unfolds.

This set of companion articles describes the two main functional components of the integrated risk-based system and their interrelationships. Herein, Part 2 provides a context for the state-of-the-art modeling of natural hazards for hurricane events. After a brief background in Section 2, the hazard modeling framework is described in Section 3. A case study analysis for Hurricane Isabel (2003) in NC is presented in Section 4.

2. BACKGROUND

The physical interactions between the hazard components (atmospheric, hydrological, and coastal oceanic) are typically captured by coupling models together via input and output file exchange, so-called loose coupling (as opposed to tight coupling by integrating models at the code level). There are several examples of such coupled systems for coastal hazards used for predicting coastal impacts from storms (e.g., tropical cyclones), both in the forecast and hindcast mode. The Coupled Ocean-Atmosphere-Wave-Sediment Transport (COAWST) modeling system (Warner, Armstrong, He, & Zambon, 2010) couples the Weather Research and Forecasting (WRF) regional meteorological model (Skamarock et al., 2008), the wind-wave model Simulating Waves Nearshore (SWAN) (Booij, Ris, & Holtuijsen, 1999), and the Regional Ocean Modeling System (ROMS) (Haidvogel et al., 2000) with the Community Sediment Transport Model (CSTM) (Warner, Butman, & Dalyander, 2008). Zambon, He, and Warner (2014) used the COAWST system to study the effects of various model coupling configurations on Hurricane Ivan (2004) impacts, finding that storm intensity predictions were very sensitive to the level of coupling between the atmosphere, ocean, and wave field models. Because researchers were primarily concerned with continental- shelf scale dynamics, precipitation and hydrological routing are not included in COAWST. Regional coastal forecasting systems using COAWST include ones for the Adriatic Sea (Russo et al., 2013) and the German Bight (Staneva, Wahle, Günther, & Stanev, 2016). There are other coupled modeling systems (e.g., Beardsley, Chen, & Xu, 2013; Blumberg, Khan, & St. John, 1999; Herrington, Bruno, & Rankin, 2000) used in the coastal regions by NOAA to provide water levels, salinity, temperature, and other water quality components. Many of these coupled model systems incorporate the riverine streamflows through USGS gauge information and do not incorporate a hydrology model to determine the rainfall-runoff caused by the storms in the inland areas, which is necessary to conduct the study presented herein. In addition, these operational forecast systems are not typically tailored for tropical cyclones.

The risk-prone NC coast has motivated several studies on simulations of hurricanes impacting the state. Peng, Xie, and Pietrafesa (2004) examined the Pamlico and Albemarle Sounds storm surge and flooding response to hypothetical hurricanes represented by a simple parametric vortex wind model (Holland, 1980), finding, unsurprisingly, that more extensive flooding occurs with more intense storm forcing. The western side of Pamlico Sound, including the Neuse River, was particularly impacted by inundation, with the inundation extent generally more sensitive to the hurricane central pressure as compared to the radius to maximum winds. Additionally, slower forward storm speed generated higher surge. Sheng, Alymov, and Paramygin (2010) simulated the Chesapeake Bay's and NC Outer Banks' response to Hurricane Isabel, using the Curvilinear-Grid Hydrodynamic 3D (CH3D) and SWAN models. They found that surge/wave coupling was important along the NC open coast, with as much as 20% of the total water level accounted for by wave setup. While these studies are indicative of the need to couple models in order to adequately capture uncertainty in simulations of coastal hazards, they do not systematically address this uncertainty.

Dresback et al. (2013) describe a storm surge and wave forecasting system that includes ADCIRC and SWAN, incorporates river discharges from an inland hydrological model, and is driven by hurricane advisory sequences from the National Hurricane Center (NHC). Hurricane winds and pressure fields are computed from the advisory file by an asymmetric vortex submodel in ADCIRC, and quantitative precipitation forecasts (QPFs) from NOAA forced the hydrology model. Hurricane Irene (2011) occurred shortly after the system's implementation, and the forecast system output over the forecast advisory sequence was compared to water-level gauges and high-water marks. Skill increased significantly as the storm approached the two-day-out point, but earlier results were still useful for early decision making.

To account for uncertainty in the evolution of hurricanes and subsequent impacts, ensembles of simulations can be used, each member of which either initially contains "small" differences in initial and/or boundary condition data or different physics options or parameter settings. This collection should represent a range of conditions that evolve differently, based on the dynamics of the models. Fundamentally, the uncertainty in predictions of complex systems arises from imperfect knowledge of the system state at a specific time and location and an imperfect description of physical process at all spatial scales. In the atmosphere, for example, observations of the weather are used to derive initial conditions that populate the entire spatial domain of a model, but these observations (1) have error and (2) are incomplete. Small differences in initial conditions can lead to large differences in the model results due to inherent nonlinearities in the system. There are many examples of recent applications of ensemble techniques in hazard modeling. Hamill, Whitaker, Kleist, Fiorino, and Benjamin (2011) showed lower hurricane track errors and better prediction skill using ensemble methods to represent different initial atmosphere states. Villarini, Luitel, Vecchi, and Ghosh (2016) demonstrate using multiple numerical models to construct ensembles of predictions of seasonal hurricane activity in the North Atlantic basin. Regarding coastal storm surge, Di Liberto, Colle, Georgas, Blumberg, and Taylor (2011) showed that a three-member ensemble using different storm surge and atmospheric models could outperform a larger atmospheric ensemble, and Brown, Souza, and Wolf (2010) and Horsburgh, Williams, Flowerdew, and Mylne (2008) showed an increased storm surge predictability using ensembles. Mel and Lionello (2014) demonstrate the connection between storm surge predictions and meteorological uncertainty with application to the Venice, Italy region.

Although unrelated to coastal U.S. waters or hurricanes, Flowerdew, Horsburgh, Wilson, and Mylne (2010) and Flowerdew, Mylne, Jones, and Titley (2013) demonstrate the need for an ensemble-based approach for storm surge predictions. They describe a prediction system for the U.K. coast that uses Met Office Global and Regional Ensemble Prediction System (MOGREPS) (Bowler, Arribas, Mylne, Robertson, & Beare, 2008) to force the CS3 storm surge model (Flather, 2000). The resulting storm surge ensemble predictions are shown to have better statistical skill compared to any one deterministic simulation. However, their focus is in larger-scale extratropical and subpolar weather systems, and thus their need for very high spatial resolution in the hazard models is somewhat less compared to simulation and prediction of tropical cyclones. Relatively high model spatial resolution for the atmospheric model ($\leq$4 km grid spacing) is typically required to realistically predict a hurricane (Davis et al., 2008), but realistic tropical cyclone structures and tracks can be obtained with grid spacings as large as 25 km (Murakami et al., 2015).

3. HAZARD MODELING FRAMEWORK

The main purpose of the hazard modeling is to provide, to the evacuation model (Davidson et al., 2020), an ensemble of predictions that represent the possible ways a hurricane might evolve. Each prediction is represented with a set of inundation and wind speed maps, one for each time step over the life cycle of the hurricane. The model suite consists of the following three state-of-the-art, physics-based numerical models: the WRF model, the Coupled Routing and Excess STorage (CREST) distributed hydrological model (Wang et al., 2011), and the ADvanced CIRCulation model for storm surge, wind wave, and tide simulation (Westerink et al., 2008). Fig. 1 shows the spatial coverage of the three hazard models used and links between the models, and Table I gives basic model information. The general workflow is as follows. Simulations of the mesoscale atmosphere (WRF) produce time-dependent fields of the dependent variables, such as QPF, sea-level atmospheric pressure, and the 10-m wind velocity at a 15-minute interval. QPF is the main input to the hydrologic model (CREST), which accumulates the precipitation and routes the inland water runoff toward the coast. The downstream river runoff is used by ADCIRC as a boundary condition on inland rivers, and the WRF wind and pressure fields force ADCIRC at the surface. ADCIRC dynamically couples the astronomical tides and wave setup forcing, and thus represents a prediction of the "total water level." This multimodel process is carried out for each ensemble member run by WRF. This collection of simulations constitutes the different storm evolutions beginning from the same initialization time.

Table I. Hazard Model Basic Configuration Information

Hazard	WRF Mesoscale Atmosphere	CREST Hydrology, River Flow	ADCIRC Storm Surge, Tides, Waves
Primary inputs	GEFS-R ensembles	WRF QPF	WRF wind/pressure CREST river flow
Horizontal resolution	36 & 12 km	250 m	20 m upriver to 50 km offshore

This modeling framework is an extension of Dresback et al. (2013) but with the following distinctions: (1) wind forcing in the former is developed using North American Mesoscale (NAM) or the NHC storm track information and parametric models, whereas herein we use wind forcing information from the WRF; (2) precipitation used the QPF from the NOAA Hydrometeorological Prediction Center in the former (in forecast mode), whereas herein the precipitation is computed by WRF; (3) the National Weather Service/Hydrology Laboratory-Research Distributed Hydrologic Model (HL-RDHM) (Koren, Reed, Smith, Zhang, & Seo, 2004) was the hydrologic model used in the former, while CREST is used herein; (4) the hydrologic model resolution in the former was set at the 4-km Hydrologic Rainfall Analysis Project (HRAP) (Reed & Maidment, 1999) grid, while CREST allows variable resolution (e.g., 250 m for the Isabel case study below); and (5) ensemble modeling of the fully coupled system is employed herein, which allows meteorological uncertainties to propagate through the system, while the former was run in deterministic mode (with the exception of forcing rivers at the boundary with a 128-member ensemble mean discharge at handoff points).

We describe the hazard modeling framework and ensemble generation in Sections 3.1 and 3.2, respectively, and the Hurricane Isabel (2003) case study in Section 4. Additionally, we draw a distinction between storm surge and inundation. We use storm surge to refer to the total water-level surface computed by the ADCIRC hydrodynamic model and referenced to a standard vertical datum

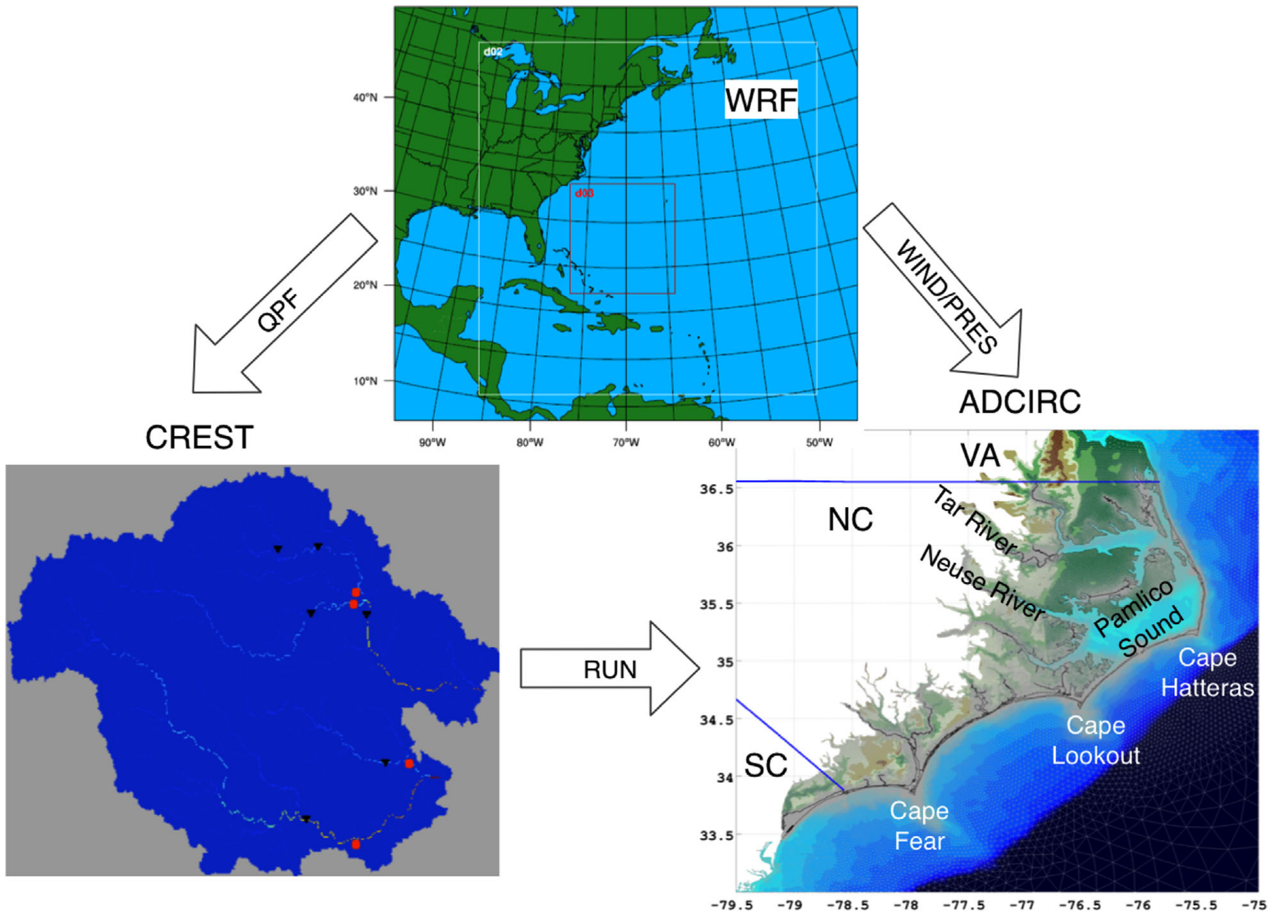

Fig. 1. Hazard model workflow. WRF computes the mesoscale atmospheric state for each hurricane scenario simulation. QPF fields are passed to CREST, which computes river flows. WRF sea-level pressure and 10-m wind velocity and CREST river flows are passed to ADCIRC, which then computes the "total water level," including astronomical tides, inland runoff, storm surge, and wave setup. Shown are the WRF spatial grid boundaries, CREST's primary river delineations along with USGS gauges (black triangles) and transfer points (red dots) in the Tar and Neuse River basins, and ADCIRC's high-resolution finite element grid for coastal NC. Main features in the area are labeled on the ADCIRC figure.

(in this case, mean sea level [MSL]). We use inundation to refer to water depth on land, meaning that the topographic elevation has been subtracted from the total water level, thus resulting in a depth over land, as referenced to the local ground elevation.

3.1. Hazard Models

3.1.1. Meteorology

Meteorological simulations for each ensemble member are computed with the numerical weather prediction model WRF (version 3.6). WRF is designed for research and operations, and is the primary regional model used by the NOAA NWS for the United States. It is highly configurable with different physics, dynamics, and parameterization schemes, a terrain-following vertical grid, and nesting of inner domains for higher resolution, including storm-following nests. WRF can use a large number of different gridded analyses or operational model forecasts for initial and boundary conditions. It solves for (among other prognostic variables) the east and north components of horizontal momentum, atmospheric pressure, and water vapor/precipitation processes, given specification of initial and boundary data. For this study, WRF is configured in a nested mode with 36 and 12 km domains. The model top vertical level is set to 50 hPa, with 45 vertical levels in total. The ensemble member selection is described in Section 3.2.

Table II. Ensemble Member Characteristics Used for Hazard Simulations

Ensemble Number	GEFS	Cumulus (36/12 km)	PBL	Microphysics	Radiation
1	P1	KF	MYJ	WSM6	RRTMG
2	P2	KF	YSU	Thompson	RRTMG
3	P3	KF	MYNN3	Goddard/3 ice	RRTM/D
4	P4	KF	BouLac	Morrison	RRTM/D
5	P5	BMJ	MYNN3	Thompson	RRTM/D
6	P6	BMJ	MYNN3	Morrison	RRTMG
7	P7	BMJ	BouLac	WSM6	RRTMG
8	P8	GD	BouLac	Goddard/3 ice	RRTMG
9	P9	GD	YSU	WSM6	RRTM/D
10	P10	GD	MYNN3	Thompson	RRTMG
11	P11	GD	MYJ	Morrison	RRTMG
12	P1	GD	MYNN3	Thompson	RRTMG
13	P1	BMJ	MYNN3	Morrison	RRTMG
14	P1	GD	YSU	Goddard/3 ice	RRTM/D
15	P1	BMJ	MYJ	WSM6	RRTM/D
16	P1	BMJ	MYNN3	Goddard/3 ice	RRTM/D
17	P2	KF	YSU	WSM6	RRTM/D
18	P4	KF	YSU	WSM6	RRTM/D
19	P5	KF	YSU	WSM6	RRTM/D
20	P7	KF	YSU	WSM6	RRTM/D
21	P8	KF	YSU	WSM6	RRTM/D
22	P10	KF	YSU	WSM6	RRTM/D

Note: GEFS is the GFS Ensemble Forecast System. Abbreviations for the options are defined in Section 3.2. RRTM/D refers to RRTM/Dudhia.

3.1.2. Hydrology

CREST is a grid-based, distributed hydrologic model developed jointly by the University of Oklahoma and the NASA SERVIR Project. CREST simulates the spatial and temporal variation of water and energy fluxes, as well as changes in storage, on a regular grid. Both global- and regional-scale applications are accommodated through subgrid representation of soil moisture storage capacity and runoff processes (using linear reservoirs or kinematic wave). CREST can be forced by gridded potential evapotranspiration and precipitation data sets from satellite-based estimates, rain gauge observations, weather radar, or combination thereof, in addition to QPFs from numerical weather prediction models. Primary water fluxes, such as infiltration and routing, are closely related to land surface characteristics. The runoff generation component and routing scheme are coupled, thus providing realistic interactions between atmospheric, surface, and subsurface water. Instead of the Sacramento Soil Moisture Accounting Model (SAC-SMA) (Burnash, Ferral, & McGuire, 1973), CREST partitions the incoming rainfall into surface runoff and infiltration using the variable infiltration capacity curve (VIC) concept, which was developed in the Xinanjiang model (Zhao, 1992; Zhao, Zhuang, Fang, Liu, & Zhang, 1980) and the VIC model (Liang, Lettenmaier, Wood, & Burges, 1994; Liang, Wood, & Lettenmaier, 1996). These latter two models are more phenomenological and do not require the large number of parameters in their calibration, whereas the SAC-SMA model does.

For the case study, the flow direction information was obtained from a digital elevation model based on the National Elevation Dataset (NED) (Gesch et al., 2002). We also used the climatological mean monthly potential evapotranspiration data (Koren et al., 2004). The STATSGO data set (Miller & White, 1998) provides several hydrologic model parameters for the model, including the soil water capacity, infiltration curve exponent, and hydraulic conductivity. Other hydrologic model parameters derived for, or used in, this model can be found in Vergara et al. (2016). For the case study herein, CREST's spatial resolution was set at 250 m, incorporating high-resolution topography, land use, and soil information maps for the target Tar/Neuse watersheds in eastern NC. The CREST model was run with precipitation input from the ensemble set shown in Table II to produce riverine flows for use in the ADCIRC model.

Data sets related to soil properties, land cover, and land use were used to form *a priori* estimates of CREST model parameters (Vergara et al., 2016). With this approach, model calibration to site-specific observations is not needed. The configured model was assessed using the Stage 4 radar quantitative precipitation estimates (the name given to the NOAA precipitation reanalysis from multisensors, e.g., gauges and radars) (Lin & Mitchell, 2005) as the surface boundary condition to CREST for Hurricane Isabel. Vergara et al. (2016) discuss the validity of this approach, particularly for ungaged locations in the watershed. Furthermore, in the current application, model uncertainty is addressed through the ensemble approach taken for the mesoscale meteorological simulations with WRF.

3.1.3. Storm Surge

The storm surge and wave simulations are conducted using the ADCIRC model, which solves a form of the shallow water wave equations, using linear, triangular finite elements in space, and a three-level finite difference discretization in time. We use ADCIRC in its two-dimensional, vertically integrated form. Land elevations and water bottom depths, frictional characteristics, land roughness lengths, and canopy cover are specified at each model node, using available digital elevation models and land cover databases, such as the USGS National Elevation Model (Gesch et al., 2002) and the National Land Cover Database (Homer, Fry, & Barnes, 2012). Spatial and temporal fields for 10-m wind velocity and atmospheric pressure are required as inputs. ADCIRC computes radiation stress gradient forcing via ADCIRC's direct coupling to the finite-element version of SWAN (Dietrich et al., 2011; Zijlema, 2010) to account for effects of wave setup on coastal water levels.

ADCIRC computes the "total" water level by simulating the coupled storm surge and tide water levels, downstream river flow and associated water levels, and contributions from wave-induced setup. The output from all three models, for each ensemble member, is then processed for input into the evacuation model. The evacuation model (Davidson et al., 2020) operates on larger geographic zones (on the order of zip code zones), so the hazard model output must be aggregated onto this coarser spatial scale.

For each simulation in the case study, the ADCIRC model is started from rest 45 days prior to ingesting the meteorological fields, with tidal boundary conditions for the M_2, N_2, S_2, K_2, K_1, O_1, P_1, and Q_1 constituents extracted from the Topex Poseidon global tidal solutions (Egbert & Erofeeva, 2002) (version 7.2) and applied at the open boundary. The tidal forcing amplitude and phase are adjusted from the equilibrium values to the specific start of the simulation, a 10-day ramp is applied, and a time step of 0.5 s is used due to the high spatial resolution. The model domain covers the North Atlantic region west of 60^{o}W, with very high spatial resolution in the NC area to support coastal inundation needed for the evacuation model (Fig. 1). The eastern boundary is the only open boundary where the tidal forcing is applied. ADCIRC's triangular finite element grid is 50 km offshore to about 20–50 m in the upper reaches of the coastal rivers (where CREST river flows are used as upstream boundary conditions). The ADCIRC grid covers the land area up to about the 15 m land contour.

3.2. Ensemble Scenario Generation

For this study, we assume that the primary uncertainty in simulating water levels and wind speeds for the evacuation model arises from imperfect knowledge of the meteorological fields used to drive the hydrological and storm surge models. This assumption is reasonable, as previous studies (Peng et al., 2004; Zhong, Li, & Zhang, 2010) have found that simulated storm surge in the Chesapeake Bay and NC estuary systems was strongly influenced by a storm's track and intensity. In neither study were parameters in the storm surge model varied, and precipitation-driven river flows were not considered. In our hazard modeling system, we have used the best available data sources for topography, bathymetry, land cover, and other fields for the hydrological and storm surge models. This will facilitate future studies on multiple sources of uncertainty and impacts on the total water-level hazard.

We represent uncertainty in hurricane predictions through two sources: (1) imperfect knowledge of the atmosphere used to initialize the meteorological model, and (2) the physics/dynamics assumptions taken when selecting various options for the meteorological model simulations. We construct an ensemble of meteorological simulations that capture this uncertainty by selecting atmospheric states and different physics options.

In general, there are several possible sources for initial and boundary data for the WRF ensemble. For contemporary, real-time operations, global ensemble simulations, such as Global Ensemble Forecast System (GEFS) (Whitaker, Hamill, Wei, Song, & Toth,

2008) or the European Centre for Medium-Range Forecasting (ECMWF) system (Molteni, Buizza, Palmer, & Petroliagis, 1996), can be used as data sources. However, their availability does not extend back before 2008. Therefore, we constructed an ensemble of meteorological simulations to reflect this uncertainty by using the 11-member GEFS Reforecast (GEFS-R) (Hamill et al., 2013) at 1-degree grid spacing for initial and boundary condition data. The reforecast is a long-running sequence (December 1984 to present) of simulations of the atmosphere that produces historical weather forecasts for research, using a recent version of the GEFS.

We constructed a 22-member ensemble of meteorological simulations to account for uncertainty by using the 11-member GEFS-R members described above. We combined these 11 members with 11 different physics option combinations to create a 22-member ensemble. Table II summarizes the permutations on WRF model physics and initial and boundary data for the ensemble. Ensemble members 1–11 correspond to the 11-member GEFS-R, and each is run with different combinations of the cumulus, planetary boundary layer, microphysics, and radiation options available in WRF. Ensemble members 12–16 use the first GEFS-R member and different physics options, and members 17–22 use the same physics options (KF, YSU, and WSM6) with six different GEFS-R ensemble members (P2, P4, P5, P7, P8, and P10). Initial and boundary conditions for WRF are extracted from the GEFS-R output data every 6 hours starting at 0000 UTC. Sea surface temperatures also have several possible sources, and for the case study, we used the 1/12th degree sea surface temperature data from the National Centers for Environmental Prediction (NCEP) for the closest 0000 UTC time to initialization.

To determine our choices of the different model-physics-based ensemble members (Table II), we conducted a review of several previous studies using WRF for modeling hurricanes (Davis et al., 2008; Fovell & Su, 2007; Fovell et al., 2016; Li & Pu, 2009; Nasrollahi et al., 2012; Nolan, Zhang, & Stern, 2009; Raju, Potty, & Mohanty, 2011; Tao et al., 2011). Most of these studies tested different combinations of two or three of the four model physics components. For the cumulus parameterizations, we used the Kain–Fritsch (KF), Betts–Miller–Janjic (BMJ), and Grell–Dvnyi (GD) schemes. The KF scheme was chosen since it has been as good as, if not better than, some of the other cumulus schemes (Biswas, Bernardet, & Dudhia, 2014). The Yonsei University (YSU), Mellor–Yamada–Janjic (MYJ), Bougeault and Lacarrere (BouLac), and Mellor–Yamada–Nakanishi–Niino level 3 (MYNN3) schemes were used for the boundary layer physics. The WRF Single-Moment 6-class (WSM6), Thompson, and Goddard (with graupel), and Morrison schemes were used for microphysics options. Fovell et al. (2016) suggested that model hurricane tracks and intensities are also sensitive to the choice of radiation physics, which interact with the cloud microphysics. They have qualitatively tested this in idealized simulations, using the rapid radiative transfer model (RRTM/Dudhia), an alternate version of the RRTM for global climate model applications (RRTMG), and the GFDL schemes, although the last of these was shown to poorly represent the interactions with deep clouds. We also used the WRF option that overrides the default formulation of the surface layer over water, which prevents the surface exchange coefficients from continuing to increase exponentially with increasing surface winds over 30 m/s.

Finally, given our ensemble-based approach that reflects the uncertainty inherent in simulating natural systems, and given the risk-based approach of the evacuation model, our hazard modeling intent is not to hindcast historical events as perfectly as possible, but rather to capture the range of outcomes, short-term trends, and likely worst-case characteristics. All three models (WRF, CREST, and ADCIRC) have been extensively validated and verified in previous studies, as reflected in the literature cited herein.

4. CASE STUDY—HURRICANE ISABEL (2003)

NC has a long and destructive history of hurricanes (Barnes, 2013)—a category 2+ hurricane is expected to impact its coast on average every 3.4 years (Vickery & Blanton, 2008), and at least tropical storm force winds about once every 2.1 years (State Climate Office of North Carolina, 2016). The low-lying coastal plain and adjacent lower piedmont area has 3.15 million residents, with the coast being a major tourist destination and eastern NC economic driver. This area is characterized by complex coastal morphology that includes the largest sheltered sound system in the United States (Pamlico and Albemarle Sounds), a thin strip of barrier islands (the Outer Banks), and large coastal rivers (Tar, Neuse, Cape Fear). In September 2003, a storm originated as a tropical wave in the eastern Atlantic Ocean and intensified to a category 5 hurricane named Isabel on September 11 (NOAA/NWS, 2004). On September 18, 2003, 17 UTC (2 p.m. local time), Hurricane Isabel made landfall as a substantially weakened but still destructive category 2 storm near Drum Inlet on the NC Outer Banks, causing about $170 million in insured property damage and three deaths in NC (Post Buckley Schuh and Jernigan, Inc., 2005). It is thus a reference storm for hazard modeling due to extensive observations of water level, winds, waves, precipitation, and river flow. The evacuation activities during the event are also well documented (Post Buckley Schuh and Jernigan, Inc., 2005)

For this Isabel case study, which also serves as the input for the case study in the companion article by Davidson et al. (2020), we applied the hazard modeling framework described in Section 3 to compute the ensemble. The case study ensemble was initialized at September 12, 2003 0000 UTC, and WRF was run for 7 days with surface variables output every 15 minutes. In Section 5, we discuss the implications of the initialization time.

4.1. Ensemble Performance and Characteristics

4.1.1. Meteorology

The WRF ensemble (Fig. 2) has a wide range of outcomes that are generally consistent with the long-range forecast prior to landfall. From this

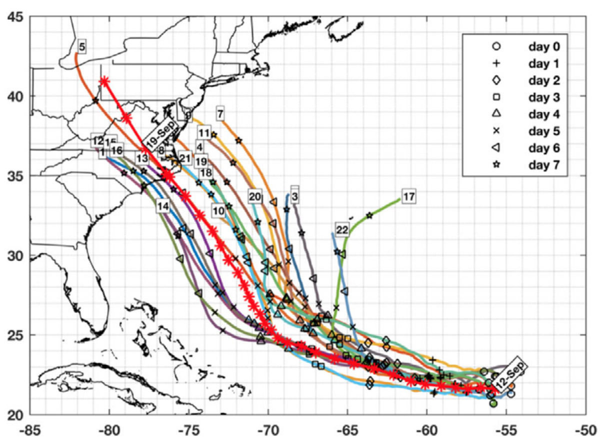

Fig. 2. Hurricane tracks for the 22-member ensemble computed for the Hurricane Isabel (2003) case study, and the best track (red) for the time period of the ensemble simulation initialized at September 12, 2003 0000 UTC, approximately 7 days prior to landfall. Daily positions of the ensemble tracks are shown with different symbols, and the ensemble member number is indicated at the end of each track.

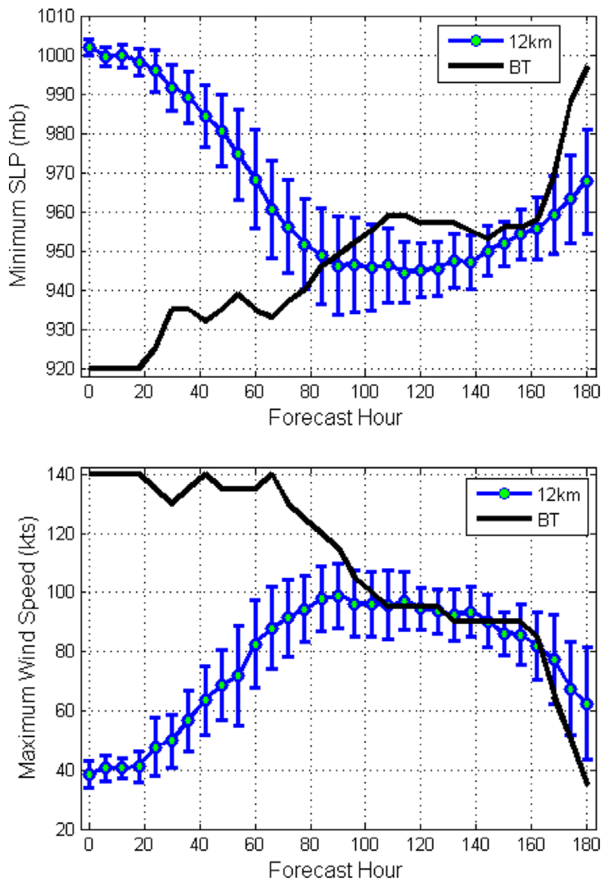

Fig. 3. Central pressure (top) (mb) and maximum wind speed (bottom) (kts) for the Hurricane Isabel ensemble initialized on September 12, 2003 0000 UTC. The observed (best track) is shown with the black line, and the 22-member ensemble mean and standard deviation are shown in blue.

offshore starting time, several members do not make landfall and five have relatively slow translation speeds and are well offshore at day 7. Six members are approaching the mid-Atlantic coast, and six have made landfall along the lower NC coast. Ensemble member (scenario) 5 follows the observed Isabel best-track path (red in Fig. 2), and makes landfall very near to the actual location near Drum Inlet, although somewhat slower and less intense than the actual hurricane. The simulated hurricane intensity, as reflected in the central pressure relative to the ambient atmospheric pressure, is shown in Fig. 3.

All ensemble members have initial pressures significantly higher that the observed hurricane's pressure, indicating a relatively weak cyclone in the GEFS reforecast at that time. However, as the hurricanes evolve, the central pressures drop and bracket the observed pressure by hour 80. Maximum sustained wind speeds (Fig. 3) show a similar behavior, with very low speeds at initialization and general agreement by hour 100. The deintensification in storm intensity (rapid increase in central pressure and decrease in maximum wind speeds) once the storm reaches land, at about hour 160, is evident in the observations (best track), but it is not as pronounced in the ensemble average. A few members that have made landfall do reflect this intensity decrease. This well-known phenomenon (Kaplan & DeMaria, 1995, 2001; Powell & Houston, 1996; Tuleya, Bender, & Kurihara, 1984; Vickery, 2005; Vickery & Twisdale, 1995) is due to the loss of thermal energy and the significant increase in surface roughness as the storm encounters land, and it is critical for models to capture because the maximum sustained wind speed on land is an important consideration in the hurricane evacuation decision-making process.

4.1.2. Hydrology

For each WRF ensemble member for Hurricane Isabel, CREST was initiated with a year-long simulation using Stage 4 radar products and then switched over to the WRF prediction of precipitation, emulating how the system would be configured in a forecasting application. As part of the data processing, the 12-km WRF ensemble members are downscaled to the 250-m CREST grid. For each CREST simulation, the QPF from WRF is used for the 7 days provided and no additional rainfall is included after the 7 days. However, the hydrologic model is run for several days after the QPF ends in order to capture the subsequent downstream routing of the runoff that occurs from this rainfall event. This additional response time is needed, as the drainage area is about 2,400 km^2 and the basins have an elongated shape. It should also be noted that the first peak seen in the streamflow results is actually a combination of the WRF-QPFs and the initial state provide by the Stage 4 derived precipitation.

The precipitation estimates for the WRF ensemble mean associated with the 7-day forecast period are shown in Fig. 4 for the western Atlantic and eastern United States and a closeup of the Tar-Pamlico and Neuse River basins for the same forecast periods. Fig. 5 shows the basin-averaged rainfall intensities for the two different watersheds for all the ensemble members, along with their mean (blue) and the Stage 4 (red) precipitation measurement. From these figures, it is evident that several of the ensemble members predict rainfall intensities that exceed the actual Stage 4 radar information. This increase in rainfall intensity is likely due to the overestimation of the hurricane intensity by WRF between hours 96 and 140 (Fig. 3), thus leading to the increase in the rainfall amounts. This is especially evident in the Neuse River basin (gray lines in Fig. 5b), where there is a threefold increase in the rainfall intensities (as compared to observed) for some of the ensemble members. However, even with these few outliers on the high side, there is an overall underestimation of the rainfall intensity when looking at all of the ensemble members for both river basins (cf. the ensemble mean [blue lines] in Fig. 5).

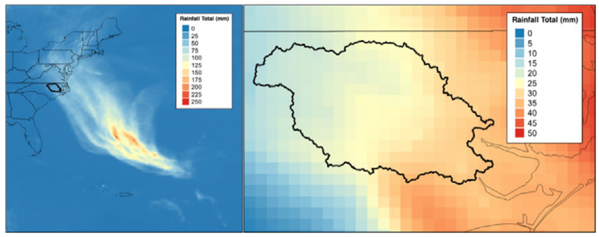

Fig. 4. (Left) Precipitation estimates (cumulative amounts, in mm) from the atmospheric model, WRF, for Hurricane Isabel for the initialization time of September 12, 2003 0000 UTC. (Right) closeup (basin outlined in black in left) of total precipitation from the same forecast period with the outline of the Tar-Pamlico and Neuse River basins. Note that the displayed color range is different in each figure.

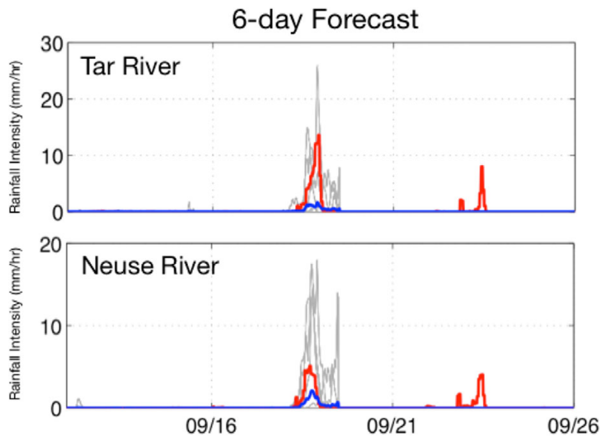

Fig. 5. Rainfall (basin-averaged) intensity (mm/hour) from the atmospheric model, WRF, for Hurricane Isabel for the Tar River basin (top) and the Neuse River basin (bottom) for the initialization time of September 12, 0000 UTC. The red line indicates the rainfall intensities from the Stage 4 radar, the gray lines provide the results from each of the 22 ensembles, and the blue line is the mean of the 22 ensemble results.

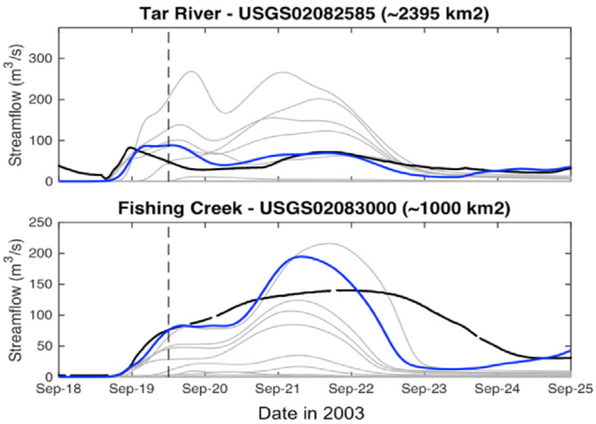

Fig. 6. Streamflow (in m^3/s) in the Tar River at the USGS stream gauge 02082585 (a) and Fishing Creek at the USGS stream gauge 02083000 (b) for Hurricane Isabel for the initialization time of September 12 0000 UTC. Gray lines are flows for each of the 22 ensemble members. Note that some members do not produce appreciable overland precipitation fields and hence no streamflow. Blue lines are modeled flows using Stage 4 radar quantitative precipitation estimates. Black lines are the observed streamflow during Hurricane Isabel. The vertical dashed line denotes the hurricane's observed landfall time.

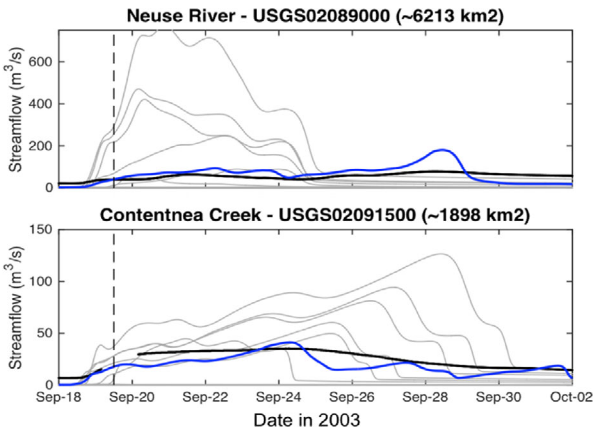

Fig. 7. Same as Fig. 6, except for the Neuse River at the USGS stream gauge 02089000 (a) and Contentnea Creek at the USGS steam gauge 02091500.

This precipitation field is then used in CREST to obtain the streamflow estimates. Ensemble members that predict the hurricane will stay far offshore produce no significant overland precipitation, and hence very small river flows. Figs. 6 and 7 show the hydrographs for each of the ensemble members, along with the hydrograph produced by the Stage 4 radar quantitative precipitation estimate, for the Tar and Neuse rivers and tributaries, respectively, included in CREST. Also shown are the observations from the USGS gauge site (black). In these results, we can see the influence of the overprediction of the rainfall intensities (gray lines in Fig. 5) with the overestimation of the streamflows. However, it should be noted that the streamflow produced by the Stage 4 radar precipitation estimate (blue line) closely follows the trend of the observations at the different gauge locations, thus validating the use of the *a priori* approach for selecting parameters, documented in Vergara et al. (2016). It is clear that an accurate hurricane track and precipitation forecast is one of the most critical elements for accurate streamflow modeling.

An important observation from these results is that the WRF precipitation estimates may vary widely in intensity and spatial distribution, but the ensemble developed from the estimates captures the general behavior of the streamflow observations from Hurricane Isabel. Thus, the ensemble methodology provides a representative probabilistic assessment for the riverine streamflows to be used by the coastal storm surge model, along with the evacuation model described in the companion article.

4.1.3. Storm Surge

We now compare the storm surge response to NOAA water-level gauge locations that were active during Hurricane Isabel (Fig. 8). This includes

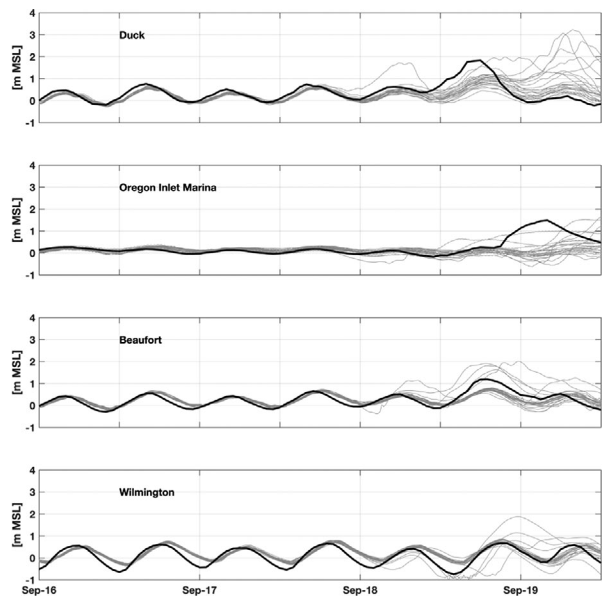

Fig. 8. Time series of observed (black) and simulated (gray) water level for Hurricane Isabel case #1, initialized at September 12 0000 UTC, at four NOAA water-level gauges in coastal NC. The first four simulation days are not shown.

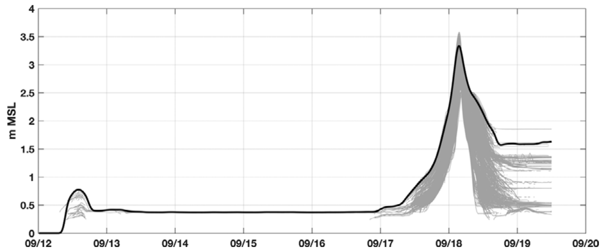

Fig. 9. Inundation time series at each ADCIRC model node (gray) in an evacuation zone in Pamlico Sound and the aggregated inundation level for that zone (black). The aggregated level is the 5% chance exceedance level given all of the wetted ADCIRC nodes at that time level. Since the evacuation zones are entirely on land, most of the ADCIRC nodes in the zone are dry during the simulation except when the storm is approaching the coast (about a day prior to landfall). The small bump at about mid-day on September 12 is due to down-river flow from precipitation used to start the CREST model.

gauges at Duck, Oregon Inlet Marina, Beaufort, and Wilmington (about 50 km up the Cape Fear River). Verified hourly water levels (surge + tide) in MSL showed peak storm surge levels of about 2 m at Duck, well to the north of the landfall location. Oregon Inlet showed a peak surge delayed by about 12 hours, which is expected due to the gauge location being well behind the barrier island. There is effectively no observed storm surge at Wilmington, which is far from the observed storm path and substantially up the Cape Fear River. Most of the ensemble members show only small surge levels, with about five showing larger and later surges than observed. Ensemble member 5 shows a peak level of 1.8 m that occurred early by about 12 hours. While the observed Hurricane Isabel showed no impacts in the Wilmington River, several ensemble members indicate that water levels above the tide level (and possible evacuation considerations) are possible. The modeled storm surges generally bracket the observed water level, although there are clearly peak arrival time differences directly related to the dispersion, intensity variation, and forward speed variation of the ensemble hurricanes, particularly noted at Duck.

4.2. Aggregation

As noted in Section 3.1.3, the final step for each ensemble simulation is to aggregate the hazard results onto the evacuation model spatial grid. Zones are considered inundated if, at any time level, at least 10% of a zone's surface area is wetted by the predicted ADCIRC water-level field. An example of this aggregation step is shown in Fig. 9, where the water-level time series for every ADCIRC node within an evacuation zone is used to construct a representative water-level curve of the 5% chance exceedance inundation level in that zone. The zone-level hazard data are used in the evacuation modeling process to create spatial maps of the hazards (inundation level and wind speeds) at each 15-minute time step. Fig. 10 shows an example of these maps for one time step.

5. DISCUSSION

The case study described above shows several interesting features that relate to the spread or dispersion of the hurricane paths and intensity within the ensemble set. Uncertainty of the hurricane path at the 7-day lead time is large enough that there is a strong threat to the middle (Beaufort) and

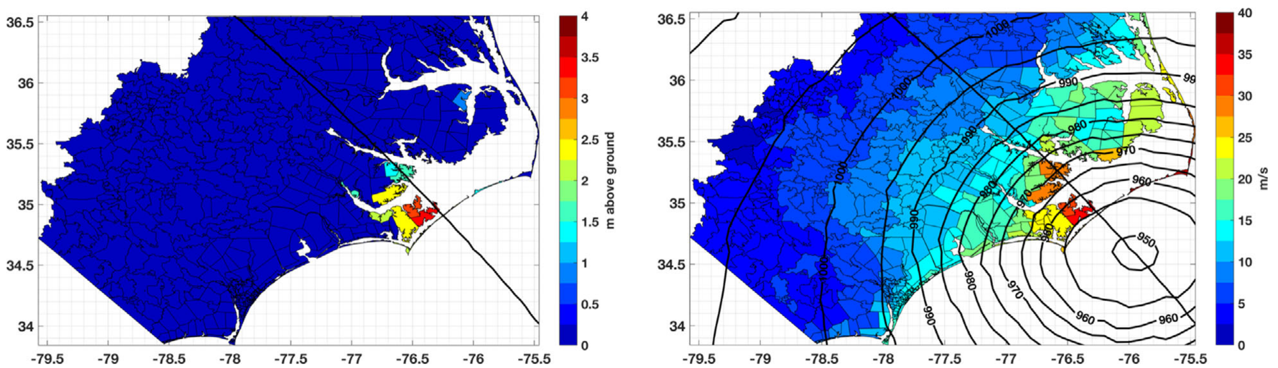

Fig. 10. Hazard maps for evacuation model zones (black outlines) at the time of the highest inundation level for ensemble member (scenario) 5. The hurricane track is shown with the black line. (a) Aggregated inundation level (5% chance exceedance level) in each evacuation model zone. (b) WRF winds in each zone, as well as contours of sea-level pressure in millibars (for reference only, not used by the evacuation model).

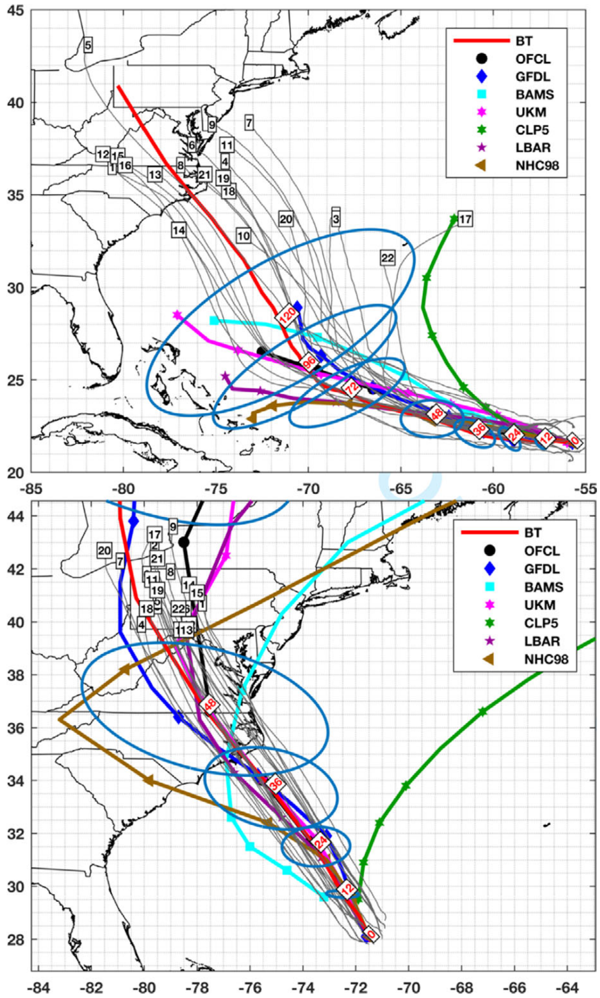

Fig. 11. Hurricane Isabel track forecasts for several operational models, for the initialization times of September 12 0000 UTC (top) and September 17 0000 UTC (bottom). Colored lines are the tracks for CLP5, GFDL, BAMS, UKM, LBAR, and NHC98, with the heavy black line being the official (OFCL) NHC forecast track. The red line is the NHC best track (BT) for Isabel. Gray lines are the 22 ensemble members for this project. Symbols on each forecast model track indicate the 0-, 12-, 24-, 36-, 48-, 72-, 96-, and 120- hour forecast positions. These hour positions are specifically labeled on the best track. Track error variance ellipses are drawn at each forecast hour position.

southern NC coast and Cape Fear River areas (Wilmington), consistent with the diversity of meteorological outcomes in this case. Our ensemble spread is qualitatively compared to the actual meteorological model forecasts for this initialization time in Fig. 11, which shows several official forecast model tracks used operationally by the NHC, as well as the NHC best track in red (determined after the storm event). The NHC forecast tracks generally follow the best track until about day 4 (96 hours), at which time the models begin to follow different trajectories. Track error variance ellipses have been drawn using the standard deviation of the differences between the forecast model track positions and the best track positions. At any specific prediction time, about two-thirds of the track positions fall within the ellipse. This is analogous to the "cone of uncertainty" used by NHC to describe likely (at the 66.6% chance level) storm eye locations. In this analysis, we have included the 5-day climatology and persistence prediction (CLP5) model (Knaff, DeMaria, Sampson, & Gross, 2003) that reflects longer term tendency of hurricane evolution toward climatology.

The evacuation model described in the companion article (Davidson et al., 2020) requires as input a single ensemble developed with one initialization time. The main benefit of an ensemble modeling approach is that it allows the evacuation model to explicitly address uncertainty in hurricane evolution and thereby determine a tree of risk-based evacuation order recommendations, each of which is conditional on how the hurricane evolves to that point in time. It thus makes the most use of the information available within each forecast. Nevertheless, as with any evacuation decision support model, when a new hazard forecast becomes available, the one described in Davidson et al. (2020) can be rerun and a new updated tree of recommendations can be generated. As with NHC forecasts, we expect that as lead time diminishes (i.e., a storm approaches shore), the uncertainty predicted by this hazard framework will generally decrease.

To illustrate the improvement in predictability as the landfall lead time decreases, we consider a second ensemble set, this one initialized at September 17, 2003 0000 UTC, 5 days later than the one in Section 4, and about 1.5 days prior to landfall. The ensemble mean minimum sea-level pressure and maximum wind speeds are shown in Fig. 12. By this time, the initial condition pressures and wind speeds in the GEFS ensembles are much more reflective of the observed pressure and generally bracket the pressure and speed by hour 30, just prior to landfall. The rapid increase in minimum pressure and decrease in wind speeds is clearly seen starting at hour 40.

The tracks for this set are shown in Fig. 11 (bottom), along with NHC forecast models and best track in red. The forecast tracks are generally centered about the best track, with a slight bias to the right. Note that the forecast model track error variance

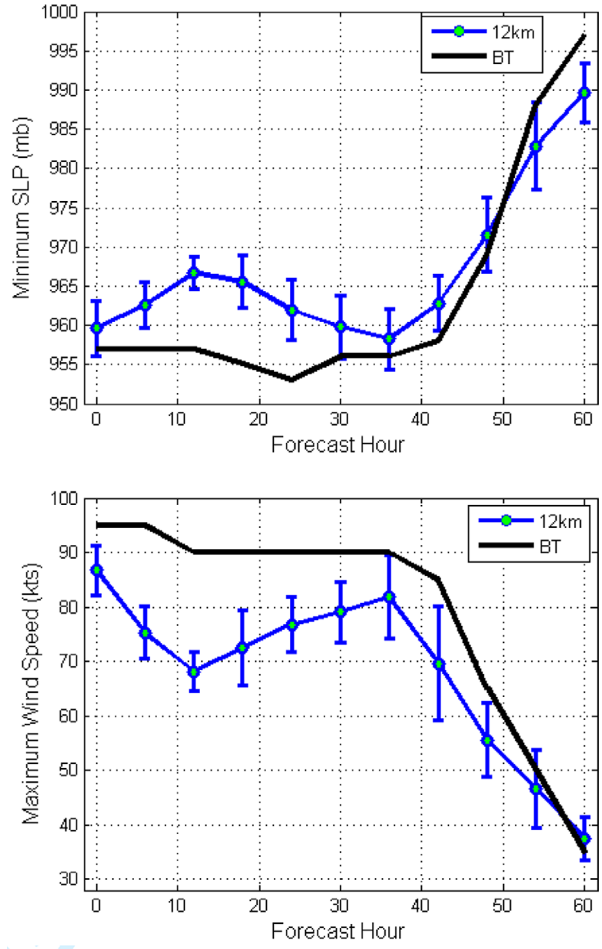

Fig. 12. Same as Fig. 3 except for the initialization time of September 17 0000 UTC. The run duration is shorter for this start time, only 2.5 days compared to 7 days at the earlier start time.

ellipses are more aligned with the mean track direction, indicating that the forward speed of the forecast storm positions is larger relative to a directional error when compared to the track dispersion from the ensemble initialized at the earlier 7-day lead time. The CLP5 climatology track is clearly influencing this error characterization in both cases, but still serves as a useful indicator of the climatological tendency of hurricane tracks to curve to the north east in the North Atlantic Basin.

The closer landfall lead time also significantly improves the WRF precipitation and subsequent hydrologic model predictions. Fig. 13 summarizes the basin-average rainfall intensities for the Tar and Neuse basins. There is an improvement in the ensemble mean capturing the rainfall intensities by comparing to the Stage 4 radar precipitation (cf. blue and red

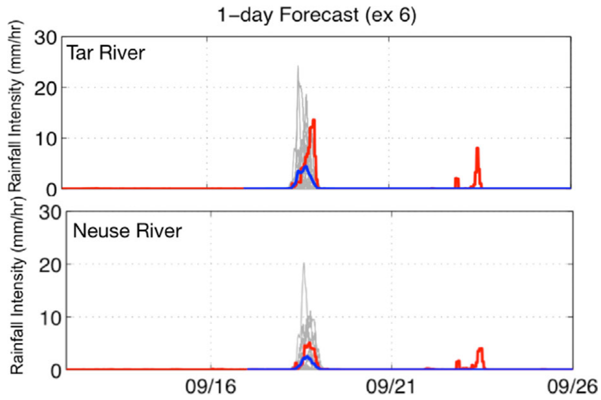

Fig. 13. Rainfall (basin-averaged) intensities (in mm/hour) from the atmospheric model, WRF, for Hurricane Isabel for the Tar River basin (a) and the Neuse River basin (b) for the initialization time of September 17 0000 UTC. The red line indicates the rainfall intensities from the Stage 4 radar, the gray lines provide the results from each of the 22 ensembles, and the blue line is the mean of the 22 ensemble results.

lines in Fig. 13), but there is still some overprediction in the rainfall intensities by a few of the ensemble members. Fig. 13 shows the streamflow for the Tar River using these ensemble precipitation estimates, along with the streamflow based on the Stage 4 radar quantitative precipitation estimates (blue line) and the USGS observed (black lines). As can be seen in Fig. 13, some ensemble members overpredict the streamflow due to the overestimation of the rainfall intensities (gray lines in Fig. 13). However, overall, we see that for many of the ensemble members, the modeled results represent observed results. Similar results (not shown) occur in the Neuse River. An exception is the shape of the hydrograph for Fishing Creek, which has an extended recession time (Fig. 13 bottom, black line after September 22); we believe that this is attributable to groundwater interflow not being captured by the hydrologic model. In particular, CREST only simulates vertical infiltration, not lateral groundwater flow. Thus, in low-lying regions with high groundwater tables, interflow can lead to a slight damping of the peak and a delayed release of water to the stream after the storm has passed.

The resulting storm surge for this ensemble is shown in Fig. 14. The threat to the northern NC coast is still substantial, and the relative timing of the storm surge peak is now more consistent with the observed peak timing, although overprediction of the peak occurs for storms that are to the right (north) of the best track. Additionally, the tracks to the left (south) of the observed track indicate that the middle

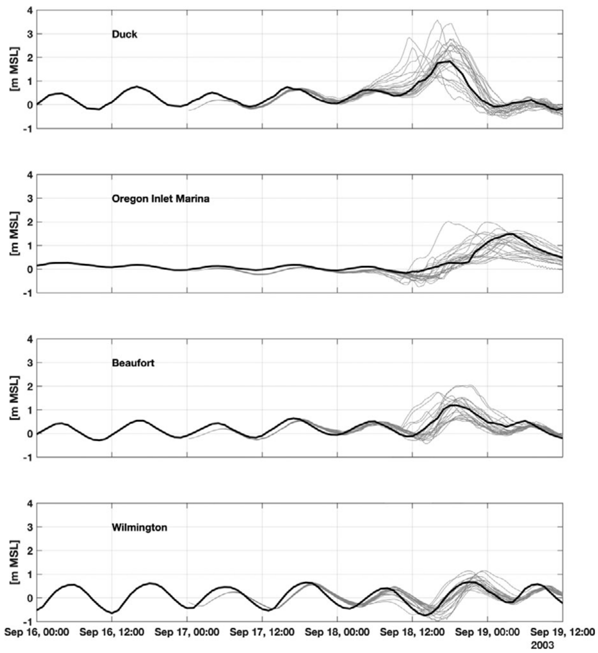

Fig. 14. Time series of observed (black) and simulated (gray) water level for Hurricane Isabel case initialized at 0000 UTC September 17, at four NOAA water-level gauges in coastal NC.

(Beaufort) and lower (Wilmington) areas are still threatened.

6. CONCLUSIONS

This article is the second of two that describes a new ensemble-based framework for hurricane evacuation modeling. This new, risk-based approach to evacuation modeling requires a probabilistic description of the hazard, which was not previously available at very high spatial and temporal resolution. It is provided here by the coupled, ensemble meteorological—hydrologic—storm surge model simulations. Herein we have focused on a description of the hazard modeling and ensemble specification and characteristics, application, and usage of which are described in Part 1 (Davidson et al., 2020) using Hurricane Isabel (2003) and its impacts in NC as a case study. The hazard simulations involve loosely coupled modeling of hurricanes and related precipitation, runoff, and storm surge. Much of the uncertainty is assumed to stem from errors in initial and boundary condition data and physics implementations in the WRF, and this uncertainty is propagated through the system via an ensemble of hurricane simulations using the WRF, CREST, and ADCIRC hazard models. The overall fidelity of the ensemble results is strongly dependent on the accuracy of the meteorological ensemble members, which improves with decreasing lead time. The ensemble envelope brackets the observed (best) track and provides "best-case" and "worst-case" end-members for the evacuation component of the modeling framework. While we have concentrated on a historical event due to the availability of information on the evacuation activities that took place, the methods are equally applicable to more contemporary events for which more accurate initialization data are available from a variety of sources. Ongoing efforts by this group include extensive validation analyses of all model components, experiments and sensitivity analyses with higher resolution (4 km) WRF ensemble members, extension of the ensemble method to include perturbations in the hydrologic and storm surge models, studies with more recent hurricane events, such as Hurricane Matthew (2016), and hardening and streamlining of the approach for use in a research forecasting mode. These activities will be reported on in future papers.

ACKNOWLEDGMENTS

The authors thank the National Science Foundation for financial support of this research under award CMMI-1331269. We also thank the anonymous reviewers for their constructive criticism, which significantly improved the article.

REFERENCES

Apivatanagul, P., Davidson, R., & Nozick, L. (2012). Bi-level optimization for risk-based regional hurricane evacuation planning. *Natural Hazards*, *60*(2), 567–588.

Barnes, J. (2013). *North Carolina's hurricane history: Updated with a decade of new storms from Isabel to Sandy*. Chapel Hill, NC: UNC Press Books.

Beardsley, R. C., Chen, C., & Xu, Q. (2013). Coastal flooding in Scituate (MA): A FVCOM study of the 27 December 2010 nor'easter. *Journal of Geophysical Research: Oceans*, *118*(11), 6030–6045.

Biswas, M. K., Bernardet, L., & Dudhia, J. (2014). Sensitivity of hurricane forecasts to cumulus parameterizations in the HWRF model. *Geophysical Research Letters*, *41*(24), 9113–9119.

Blumberg, A. F., Khan, L. A., & St. John, J. P. (1999). Three-dimensional hydrodynamic model of New York Harbor region. *Journal of Hydraulic Engineering*, *125*(8), 799–816.

Booij, N., Ris, R., & Holtuijsen, L. (1999). A third-generation wave model for coastal regions, Part I: Model description and validation. *Journal of Geophysical Research*, *104*(C4), 7649–7666.

Bowler, N. E., Arribas, A., Mylne, K. R., Robertson, K. B., & Beare, S. E. (2008). The MOGREPS short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, *134*(632), 703–722.

Brown, J. M., Souza, A. J., & Wolf, J. (APR 2010). Surge modelling in the eastern Irish Sea: Present and future storm impact. *Ocean Dynamics*, *60*(2), 227–236. https://doi.org/10.1007/s10236-009-0248-8.

Burnash, R., Ferral, R., & McGuire, R. (1973). *A generalized streamflow simulation system—Conceptual modeling for digital computers*. Technical report, Joint Federal and State River Forecast Center.

Davidson, R. A., Nozick, L. K., Wachtendorf, T., Blanton, B., Colle, B., Kolar, R. L., … Leonardo, N. (2020). An Integrated Scenario Ensemble?Based Framework for Hurricane Evacuation Modeling: Part 1—Decision Support System. *Risk Analysis*, *40*(1), 97–116.

Davis, C., Wang, W., Chen, S., Chen, Y., Corbosiero, K., DeMaria, M., … Xiao, Q. (2008). Prediction of landfalling hurricanes with the advanced hurricane WRF model. *Monthly Weather Review*, *136*(6), 1990–2005.

Dietrich, J. C., Westerink, J. J., Kennedy, A. B., Smith, J. M., Jensen, R. E., Zijlema, M., … Cobell, Z. (2011). Hurricane Gustav (2008) waves and storm surge: Hindcast, synoptic analysis, and validation in southern Louisiana. *Monthly Weather Review*, *139*(8), 2488–2522. https://doi.org/10.1175/2011MWR3611.1.

Di Liberto, T., Colle, B. A., Georgas, N., Blumberg, A. F., & Taylor, A. A. (2011). Verification of a multimodel storm surge ensemble around New York City and Long Island for the cool season. *Weather and Forecasting*, *26*(6), 922–939.

Dresback, K., Fleming, J., Blanton, B., Kaiser, C., Gourley, J., Tromble, E., & Nemunaitis-Monroe, K. (2013). Skill assessment of a real-time forecast system utilizing a coupled hydrologic and coastal hydrodynamic model during Hurricane Irene (2011). *Continental Shelf Research*, *71*, 78–94.

Egbert, G. D., & Erofeeva, S. Y. (2002). Efficient inverse modeling of barotropic ocean tides. *Journal of Atmospheric and Oceanic Technology*, *19*(2), 183–204.

Flather, R. A. (2000). Existing operational oceanography. *Coastal Engineering*, *41*(1–3), 13–40.

Flowerdew, J., Horsburgh, K., Wilson, C., & Mylne, K. (2010) Development and evaluation of an ensemble forecasting system for coastal storm surges. *Quarterly Journal of the Royal Meteorological Society*, *136*(651), 1444–1456. https://doi.org/10.1002/qj.648.

Flowerdew, J., Mylne, K., Jones, C., & Titley, H. (2013). Extending the forecast range of the UK storm surge ensemble. *Quarterly Journal of the Royal Meteorological Society*, *139*(670), 184–197.

Fovell, R.G., Y.P. Bu, K.L. Corbosiero, W. Tung, Y. Cao, H. Kuo, L. Hsu, and H. Su, 2016. Influence of Cloud Microphysics and Radiation on Tropical Cyclone Structure and Motion. *Meteorological Monographs*, *56*, 11.1–11.27, https://doi.org/10.1175/AMSMONOGRAPHS-D-15-0006.1.

Fovell, R. G., & Su, H. (2007). Impact of cloud microphysics on hurricane track forecasts. *Geophysical Research Letters*, *34*(24). https://doi.org/10.1029/2007GL031723.

Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., & Tyler, D. (2002). The National Elevation Dataset. *Photogrammetric Engineering and Remote Sensing*, *68*(1), 5–32.

Haidvogel, D. B., Arango, H. G., Hedstrom, K., Beckmann, A., Malanotte-Rizzoli, P., & Shchepetkin, A. F. (2000). Model evaluation experiments in the North Atlantic Basin: Simulations in nonlinear terrain-following coordinates. *Dynamics of Atmospheres and Oceans*, *32*(3–4), 239–281.

Hamill, T. M., Bates, G. T., Whitaker, J. S., Murray, D. R., Fiorino, M., Galarneau Jr, T. J., … Lapenta, W. (2013). NOAA's second-generation global medium-range ensemble reforecast dataset. *Bulletin of the American Meteorological Society*, *94*(10), 1553–1565.

Hamill, T. M., Whitaker, J. S., Kleist, D. T., Fiorino, M., & Benjamin, S. G. (2011). Predictions of 2010's tropical cyclones using the GFS and ensemble-based data assimilation methods. *Monthly Weather Review*, *139*(10), 3243–3247.

Herrington, T., Bruno, M., & Rankin, K. (2000). The New Jersey coastal monitoring network: A real-time coastal observation system. *Journal of Marine Environmental Engineering*, *6*(1), 69–82.

Holland, G. J. (1980). An analytic model of the wind and pressure profiles in hurricanes. *Monthly Weather Review*, *108*(8), 1212–1218. https://doi.org/10.1175/1520-0493(1980)108<1212:AAMOTW>2.0.CO;2.

Homer, C. H., Fry, J. A., & Barnes, C. A. (2012). *The National Land Cover Database*. Technical report, US Geological Survey.

Horsburgh, K., Williams, J., Flowerdew, J., & Mylne, K. (2008). Aspects of operational forecast model skill during an extreme storm surge event. *Journal of Flood Risk Management*, *1*(4), 213–221.

HURREVAC. (2016). *2016 documentation and user's manual: HURREVAC for 2016 season*. Technical Report version 1.6.1.

Jelesnianski, C., Chen, J., & Shaffer, W. A. (1992). *SLOSH: Sea, Lake, and Overland Surges from Hurricanes*. NOAA Technical Report NWS 48, U.S. Department of Commerce.

Kaplan, J., & DeMaria, M. (2001). On the decay of tropical cyclone winds after landfall in the New England area. *Journal of Applied Meteorology*, *40*(2), 280–286.

Kaplan, J., & DeMaria, M. (1995). A simple empirical model for predicting the decay of tropical cyclone winds after landfall. *Journal of Applied Meteorology*, *34*(11), 2499–2512.

Knaff, J. A., DeMaria, M., Sampson, C. R., & Gross, J. M. (2003). Statistical, 5-day tropical cyclone intensity forecasts derived from climatology and persistence. *Weather and Forecasting*, *18*(1), 80–92.

Koren, V., Reed, S., Smith, M., Zhang, Z., & Seo, D. J. (2004). Hydrology Laboratory Research Modeling System (HL-RMS) of the U.S. National Weather Service. *Journal of Hydrology*, *291*, 297–318.

Li, A., Xu, N., Nozick, L., & Davidson, R. (2011). Bilevel optimization for integrated shelter location analysis and transportation planning for hurricane events. *Journal of Infrastructure Systems*, *17*(4), 184–192.

Li, X., & Pu, Z. (2009). Sensitivity of numerical simulations of the early rapid intensification of Hurricane Emily to cumulus parameterization schemes in different model horizontal resolutions. *Journal of the Meteorological Society of Japan Series II*, *87*(3), 403–421.

Liang, X., Wood, E. F., & Lettenmaier, D. P. (1996). Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Global and Planetary Change*, *13*(1–4), 195–206.

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres*, *99*(D7), 14415–14428.

Lin, Y., & Mitchell, K. E. (2005). 1.2 the NCEP stage II/IV hourly precipitation analyses: Development and applications. In *19th Conference on Hydrology*. Citeseer.

Mel, R., & Lionello, P. (2014). Storm surge ensemble prediction for the city of Venice. *Weather and Forecasting*, *29*(4), 1044–1057.

Miller, D. A., & White, R. A. (1998). A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling. *Earth Interactions*, *2*(2), 1–26.

Molteni, F., Buizza, R., Palmer, T. N., & Petroliagis, T. (1996). The ECMWF ensemble prediction system: Methodology and validation. *Quarterly Journal of the Royal Meteorological Society*, *122*(529), 73–119. https://doi.org/10.1002/qj.49712252905.

Murakami, H., Vecchi, G. A., Underwood, S., Delworth, T. L., Wittenberg, A. T., Anderson, W. G., … Lin, S.-J. (2015). Simulation and prediction of category 4 and 5 hurricanes in the high-resolution GFDL HiFLOR coupled climate model. *Journal of Climate*, *28*(23), 9058–9079.

Nasrollahi, N., AghaKouchak, A., Li, J., Gao, X., Hsu, K., & Sorooshian, S. (2012). Assessing the impacts of different WRF precipitation physics in hurricane simulations. *Weather and Forecasting*, *27*(4), 1003–1016.

NOAA/NWS. (2004). Hurricane Isabel — September 18–19, 2003. Service Assessment.

Nolan, D. S., Zhang, J. A., & Stern, D. P. (2009). Evaluation of planetary boundary layer parameterizations in tropical cyclones by comparison of in situ observations and high-resolution simulations of Hurricane Isabel (2003). Part I: Initialization, maximum winds, and the outer-core boundary layer. *Monthly Weather Review*, *137*(11), 3651–3674.

Peng, M., Xie, L., & Pietrafesa, L. J. (2004). A numerical study of storm surge and inundation in the Croatan-Albemarle-Pamlico estuary system. *Estuarine, Coastal and Shelf Science*, *59*(1), 121–137.

Post Buckley Schuh and Jernigan, Inc. (March 2005). Hurricane Isabel Assessment: *Review of hurricane evacuation study products and other aspects of the National Hurricane Mitigation and Preparedness Program (NHMPP) in the context of the Hurricane Isabel response*. Technical Report, PBS&J, Tallahassee, FL.

Powell, M. D., & Houston, S. H. (1996). Hurricane Andrew's landfall in South Florida. Part II: Surface wind fields and potential real-time applications. *Weather and Forecasting*, *11*(3), 329–349.

Raju, P., Potty, J., & Mohanty, U. (2011). Sensitivity of physical parameterizations on prediction of tropical cyclone Nargis over the Bay of Bengal using WRF model. *Meteorology and Atmospheric Physics*, *113*(3–4), 125.

Reed, S. M., & Maidment, D. R. (1999). Coordinate transformations for using NEXRAD data in GIS-based hydrologic modeling. *Journal of Hydrologic Engineering*, *4*(2), 174–182.

Russo, A., Coluccelli, A., Carniel, S., Benetazzo, A., Valentini, A., Paccagnella, T., … Bortoluzzi, G. (2013). Operational models hierarchy for short term marine predictions: The Adriatic Sea example. In *OCEANS - Bergen, 2013 MTS/IEEE*, pp. 1–6. https://doi.org/10.1109/OCEANS-Bergen.2013.6608139.

Sheng, Y. P., Alymov, V., & Paramygin, V. A. (2010). Simulation of storm surge, wave, currents, and inundation in the Outer Banks and Chesapeake Bay during Hurricane Isabel in 2003: The importance of waves. *Journal of Geophysical Research: Oceans*, *115*(C4). https://doi.org/10.1029/2009JC005402.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., & Powers, J. G. (June 2008). *A description of the advanced research WRF version 3*. NCAR Technical Note NCAR/TN–475+STR, Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, CO.

Staneva, J., Wahle, K., Günther, H., & Stanev, E. (2016). Coupling of wave and circulation models in coastal–ocean predicting systems: A case study for the German Bight. *Ocean Science*, *12*(3), 797–806.

State Climate Office of North Carolina. (2016). http://climate.ncsu.edu/climate/hurricanes/statistics.php?state=nc.

Tao, W.-K., Shi, J. J., Chen, S. S., Lang, S., Lin, P.-L., Hong, S.-Y., … Hou, A. (2011). The impact of microphysical schemes on hurricane intensity and track. *Asia-Pacific Journal of Atmospheric Sciences*, *47*(1), 1–16.

Tuleya, R. E., Bender, M. A., & Kurihara, Y. (1984). A simulation study of the landfall of tropical cyclones. *Monthly Weather Review*, *112*(1), 124–136.

Vergara, H., Kirstetter, P.-E., Gourley, J. J., Flamig, Z. L., Hong, Y., Arthur, A., & Kolar, R. (2016). Estimating a-priori kinematic wave model parameters based on regionalization for flash flood forecasting in the conterminous United States. *Journal of Hydrology*, *541*, 421–433.

Vickery, P. J. (2005). Simple empirical models for estimating the increase in the central pressure of tropical cyclones after landfall along the coastline of the United States. *Journal of Applied Meteorology*, *44*(12), 1807–1826. https://doi.org/10.1175/JAM2310.1.

Vickery, P., & Blanton, B. (September 2008). *North Carolina Coastal Flood Analysis System: Hurricane parameter development*. Technical Report TR-08-06, RENCI, North Carolina.

Vickery, P. J., & Twisdale, L. A. (1995). Wind-field and filling models for hurricane wind-speed predictions. *Journal of Structural Engineering*, *121*(11), 1700–1709.

Villarini, G., Luitel, B., Vecchi, G. A., & Ghosh, J. (2016). Multi-model ensemble forecasting of North Atlantic tropical cyclone activity. *Climate Dynamics*, 1–17. https://link.springer.com/article/10.1007%2Fs00382-016-3369-z.

Wang, J., Hong, Y., Li, L., Gourley, J. J., Khan, S. I., Yilmaz, K. K., … Irwn, D. (2011). The coupled routing and excess storage (CREST) distributed hydrological model. *Hydrological Sciences Journal*, *56*(1), 84–98.

Warner, J. C., Armstrong, B., He, R., & Zambon, J. B. (2010). Development of a coupled ocean–atmosphere–wave–sediment transport (COAWST) modeling system. *Ocean Modelling*, *35*(3), 230–244.

Warner, J. C., Butman, B., & Dalyander, P. S. (2008). Storm-driven sediment transport in Massachusetts Bay. *Continental Shelf Research*, *28*(2), 257–282.

Westerink, J., Luettich, R., Feyen, J., Atkinson, J., Dawson, C., Roberts, H., … Pourtaheri, H. (2008). A basin- to channel-scale unstructured grid hurricane storm surge model applied to southern Louisiana. *Monthly Weather Review*, *136*, 833–864.

Whitaker, J. S., Hamill, T. M., Wei, X., Song, Y., & Toth, Z. (2008). Ensemble data assimilation with the NCEP global forecast system. *Monthly Weather Review*, *136*(2), 463–482.

Whitehead, J. C. (2003). One million dollars per mile? The opportunity costs of hurricane evacuation. *Ocean and Coastal Management*, *46*(11–12), 1069–1083. https://doi.org/10.1016/j.ocecoaman.2003.11.001.

Zambon, J. B., He, R., & Warner, J. C. (2014). Investigation of Hurricane Ivan using the coupled ocean–atmosphere–wave–sediment transport (COAWST) model. *Ocean Dynamics*, *64*(11), 1535–1554.

Zhao, R.-J. (1992). The Xinanjiang model applied in China. *Journal of Hydrology*, *135*(1), 371–381. https://doi.org/10.1016/0022-1694(92)90096-E.

Zhao, R. J., Zhuang, Y.-L., Fang, L.-R., Liu, X.-R. & Zhang, Q.-S. (1980). The Xinanjiang model. In *Hydrological Forecasting*, IAHS Publication No. 129, pages 351–356, Int. Assoc. Sci. Hydrol. Press, Wallingford, UK.

Zhong, L., Li, M., & Zhang, D.-L. (2010). How do uncertainties in hurricane model forecasts affect storm surge predictions in a semi-enclosed bay? *Estuarine, Coastal and Shelf Science*, *90*(2), 61–72.

Zijlema, M. (2010). Computation of wind-wave spectra in coastal waters with SWAN on unstructured grids. *Coastal Engineering*, *57*(3), 267–277.